REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
27-07-2018	Final Report	1-Jun-2015 - 31-Jan-2018

4. TITLE AND SUBTITLE

Final Report: Statistical Structural Health Monitoring in the Presence of Environmental Variability and Uncertainty

5a. CONTRACT NUMBER
W911NF-15-1-0172

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

Harvard University
Office for Sponsored Programs
1033 Massachusetts Ave 5th Floor
Cambridge, MA 02138 -5369

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
65511-MA.11

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Luke Bornn
UU	UU	UU	UU		19b. TELEPHONE NUMBER 617-496-9259

RPPR Final Report
as of 01-Aug-2018

Agency Code:

Proposal Number: 65511MA **Agreement Number: W911NF-15-1-0172**
INVESTIGATOR(S):

Name: Luke Bornn
Email: lbornn@sfu.ca
Phone Number: 6174969259
Principal: Y

Organization: **Harvard University**
Address: Office for Sponsored Programs, Cambridge, MA 021385369
Country: USA
DUNS Number: 082359691 EIN: 042103580N
Report Date: 30-Apr-2016 Date Received: 27-Jul-2018
Final Report for Period Beginning 01-Jun-2015 and Ending 31-Jan-2018
Title: Statistical Structural Health Monitoring in the Presence of Environmental Variability and Uncertainty
Begin Performance Period: 01-Jun-2015 **End Performance Period:** 31-Jul-2018
Report Term: 0-Other
Submitted By: Luke Bornn Email: lbornn@sfu.ca
Phone: (617) 496-9259
Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: **STEM Participants:**

Major Goals: The goal of this projectl is to develop statistical methodologies to tackle the huge data and diverse environmental variability facing the field of structural health monitoring. The primary objectives are:
1. Design a statistical framework to handle the massive data generated by SHM systems
2. Develop tools to explicate the diverse environmental conditions experienced by SHM systems while maintaining fidelity of damage detection
3. Implement and test the proposed methodology on real-life structures

Accomplishments: I would conclude that we completed 2.5 out of the 3 primary goals. The remaining .5 is currently being taken up by colleagues at Los Alamos National Labs who are working on real life SHM systems on which our methodology is being tested.

1. Design a statistical framework to handle the massive data generated by SHM systems
As inexpensive sensors make vast SHM networks feasible across a wide range of engineering structures, the resulting data generated by such systems is massive, requiring methods which naturally filter and make use of the information contained within this big data. To model the full system in networks with hundreds or thousands of vibration sensors is not feasible due to the huge number of parameters requiring estimation. By exploiting the interconnectivity of the engineering structure, we can induce parsimonious models which scale naturally with growing networks. This led to some initial work (http://www.lukebornn.com/papers/bornn_mssp_2016.pdf) which uses Bayesian graphical models to understand and model structural engineering systems.

2. Develop tools to explicate the diverse environmental conditions experienced by SHM systems while maintaining fidelity of damage detection
SHM systems implemented in the real world encounter tremendous environmental and operational variability. As an example, a vibration-based SHM system employed in large-scale scientific infrastructure such as telescopes would encounter different temperatures, solar radiation levels, precipitation, internal settings, and usages. All of these different conditions might impact the telescope's natural resonant frequencies. Specifically, a telescope changes physically as zoom levels change and as a result resonant frequencies change as well. As such, if one trains the model in one zoom, then tests on another zoom, it might flag the change as being a result of damage, rather than natural usage variability. Alternatively, one might train the model across various environmental conditions, but the increased variability reduces the signal to noise ratio, making damage detection more difficult (Farrar and Lieven, 2007). As such, the models considered here must alter as the environment changes, adapting to different conditions to maintain the fidelity of damage detection.
The resolution of this task is a journal article published in Structural Health Monitoring, which explores the use of

factorial switching linear dynamical systems. These systems, which fit within the Bayesian graphical modeling framework, allow us to account for environmental and operational variability from both observed and unobserved sources. In addition, several related works were developed, as listed in "Dissemination"

Training Opportunities: The primary training opportunities coming from the reporting period were the training of three Harvard statistics graduate students and one postdoctoral fellow. These trainees gained exposure not only to the statistical methods proposed in our research, but also the underlying engineering problems. Thus, they not only advanced their expertise in the areas of statistics and machine learning, but also developed a skill base in a fast-growing application area with important real-world implications.

Results Dissemination: The primary dissemination during this period has been the submission of several manuscripts. These are:

• Gerber, M., Bornn, L. (2018) Convergence Results for a Class of Time-Varying Simulated Annealing Algorithms. Stochastic Processes and their Applications. Vol. 128, 1073-1094.
• Ward, P., Tankovich, M., Ramsden, J., Drust, B., Bornn, L. (2018) Volume and Intensity are Important Training Related Factors in Injury Incidence in American Football Athletes. Sloan Sports Analytics Conference 2018 (Finalist).
• Fernandez, J., Bornn, L. (2018) Wide Open Spaces: A Statistical Technique for Measuring Space Creation in Professional Soccer Sloan Sports Analytics Conference 2018 (Finalist).
• Liu, A., Wang, L., Bornn, L., and Farrar, C. (2018) Robust Structural Health Monitoring Under Environmental and Operational Uncertainty with Switching State-Space Autoregressive Models. To appear in Structural Health Monitoring.
• Wu, S., Bornn, L. (2018) Modeling Offensive Player Movement in Professional Basketball. To appear in The American Statistician.
• Franks, A., D'Amour A., †Cervone, D., Bornn, L. (2017) Meta-Analytics: Tools for Understand- ing the Statistical Properties of Sports Metrics. The Journal of Quantitative Analysis in Sports. Vol. 12, 151-165.
• van Bommel, M., Bornn, L. (2017) Adjusting for Scorekeeper Bias in NBA Box Scores. Data Mining and Knowledge Discovery. Vol. 31, 1622-1642.
• Gerber, M., Bornn, L. (2017) Improving Simulated Annealing through Derandomization. The Journal of Global Optimization. Vol. 68, 189-217.
• Bornn, L., Pillai, N., Smith, A., Woodard., D. (2017) The Use of a Single Pseudo-Sample in Approximate Bayesian Computation. Statistics and Computing. Vol. 27, 583-590.
• Miller, A., Bornn, L. (2017) Possession Sketches: Mapping NBA Strategies. Sloan Sports Analytics Conference 2017 (Finalist).
†Indicates students and other HQP
• Antonelli, J., Cefalu, M., Bornn, L. (2016) The Positive Effects of Population Based Preferential Sampling in Environmental Epidemiology. Biostatistics. Vol. 17, 764-778.
• Cervone, D., D'Amour, A., Bornn, L., Goldsberry, K. (2016) A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes. Journal of the American Statistical Association. Vol. 111, 585-599.
• Bornn, L., Farrar, C., Higdon, D., Murphy, K. (2016) Modeling and Diagnosis of Structural Systems through Sparse Dynamic Graphical Models. Mechanical Systems and Signal Processing. Vol. 74, 133143.
• Bojinov, I., Bornn, L. (2016) The Pressing Game: Optimal Defensive Disruption in Soccer. Sloan Sports Analytics Conference 2016 (Finalist).
• Chen, Y., Bornn, L., De Freitas, N., Eskelin, M., Fang, J., Welling, M. (2016) Herded Gibbs Sampling. Journal of Machine Learning Research. Vol. 17, 1-29.

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI
Participant: Luke Bornn
Person Months Worked: 4.00 **Funding Support:**

Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Reza Solgi
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Lazhi Wang
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Undergraduate Student
Participant: Anthony Liu
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

ARTICLES:

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 2-Awaiting Publicat
Journal: Statistics and Computing
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published: 9/13/16 3:32PM
Publication Location:
Article Title: The Use of a Single Pseudo-Sample in Approximate Bayesian Computation
Authors: Luke Bornn, Natesh Pillai, Aaron Smith, Dawn Woodard
Keywords: monte carlo, approximate bayesian computation
Abstract: We analyze the computational efficiency of approximate Bayesian computation (ABC), which approximates a likelihood function by drawing pseudo-samples from the associated model. For the rejec- tion sampling version of ABC, it is known that multiple pseudo-samples cannot substantially increase (and can substantially decrease) the efficiency of the algorithm as compared to employing a high-variance estimate based on a single pseudo-sample. We show that this conclusion also holds for a Markov chain Monte Carlo version of ABC, implying that it is unnecessary to tune the number of pseudo-samples used in ABC-MCMC. This conclusion is in contrast to particle MCMC methods, for which increasing the number of particles can provide large gains in computational efficiency.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 01-Aug-2018

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 4-Under Review
Journal: Unpublished
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published: 9/13/16 3:37PM
Publication Location:
Article Title: Nonparametric Hierarchical Bayesian Quantiles
Authors: Luke Bornn, Neil Shephard, Reza Solgi
Keywords: Bayesian Nonparametrics, Quantiles
Abstract: Here we develop a method for performing nonparametric Bayesian inference on quantiles. Relying on geometric measure theory and employing a Hausdorff base measure, we are able to specify meaningful priors for the quantile while treating the distribution of the data otherwise nonparametrically. We further extend the method to a hierarchical model for quantiles of subpopulations, linking subgroups together solely through their quantiles. Our approach is computationally straightforward, allowing for censored and noisy data. We demonstrate the proposed methodology on simulated data and an applied problem from sports statistics, where it is observed to stabilize and improve inference and prediction.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 4-Under Review
Journal: Unpublished
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published: 9/13/16 3:41PM
Publication Location:
Article Title: Moment Conditions and Bayesian Nonparametrics
Authors: Luke Bornn, Neil Shephard, Reza Solgi
Keywords: Bayesian nonparametrics
Abstract: Models phrased though moment conditions are central to much of modern inference. Here these moment conditions are embedded within a nonparametric Bayesian setup. Handling such a model is not probabilistically straightforward as the posterior has support on a manifold. We solve the relevant issues, building new probability and computational tools using Hausdorff measures to analyze them on real and simulated data. These new methods which involve simulating on a manifold can be applied widely, including providing Bayesian analysis of quasi- likelihoods, linear and nonlinear regression, missing data and hierarchical models.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 01-Aug-2018

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 2-Awaiting Publicat
Journal: Journal of Global Optimization
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published: 9/13/16 3:44PM
Publication Location:
Article Title: Improving Simulated Annealing through Derandomization
Authors: Mathieu Gerber, Luke Bornn
Keywords: Global optimization; Quasi-Monte Carlo; Randomized quasi- Monte Carlo; Simulated annealing; Threshold accepting
Abstract: We propose and study a version of simulated annealing (SA) on continu- ous state spaces based on (t, s)R -sequences. The parameter R ? N ? regulates the degree of randomness of the input sequence, with the case R = 0 corres- ponding to IID uniform random numbers and the limiting case R = ? to (t, s)-sequences. Our main result, obtained for rectangular domains, shows that the resulting optimization method, which we refer to as QMC-SA, con- verges almost surely to the global optimum of the objective function ? for any R ? N. When ? is univariate, we are in addition able to show that the completely deterministic version of QMC-SA is convergent.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 0-Other
Journal: Unpublished
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published: 9/13/16 3:48PM
Publication Location:
Article Title: FastGP: An R Package for Gaussian Processes
Authors: Giri Gopalan, Luke Bornn
Keywords: Gaussian Processes
Abstract: Despite their promise and ubiquity, Gaussian processes (GPs) can be difficult to use in practice due to the computational impediments of fitting and sampling from them. Here we discuss a short R package for efficient multivariate normal functions which uses the Rcpp and RcppEigen packages at its core. GPs have properties that allow standard functions to be sped up; as an example we include functionality for Toeplitz matrices whose inverse can be computed in O(n2) time with methods due to Trench and Durbin (Golub & Van Loan 1996), which is particularly apt when time points (or spatial locations) of a Gaussian process are evenly spaced, since the associated covariance matrix is Toeplitz in this case. Additionally, we include functionality to sample from a latent variable Gaussian process model with elliptical slice sampling (Murray, Adams, & MacKay 2010).
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 01-Aug-2018

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 1-Published
Journal: Journal of Quantitative Analysis in Sports
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: Meta-Analytics: Tools for Understanding the Statistical Properties of Sports Metrics
Authors: Alexander Franks, Alexander D'Amour, Daniel Cervone and Luke Bornn
Keywords: Statistics
Abstract: In sports, there is a constant effort to improve metrics which assess player abil- ity, but there has been almost no effort to quantify and compare existing metrics. Any individual making a management, coaching, or gambling decision is quickly over- whelmed with hundreds of statistics. We address this problem by proposing a set of "meta-metrics" which can be used to identify the metrics that provide the most unique, reliable, and useful information for decision-makers. Specifically, we develop methods to evalute metrics based on three criteria: 1) stability: does the metric measure the same thing over time 2) discrimination: does the metric differentiate between players and 3) independence: does the metric provide new information? Our methods are easy to implement and widely applicable so they should be of interest to the broader sports community. We demonstrate our methods in analyses of both NBA and NHL metrics. Our results indicate the most reliable metrics and highlight how they sh
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 1-Published
Journal: Biostatistics
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: The positive effects of population-based preferential sampling in environmental epidemiology
Authors: Joseph Antonelli, Matthew Cefalu, Luke Bornn
Keywords: Statistics, Preferential Sampling
Abstract: In environmental epidemiology, exposures are not always available at subject locations and must be predicted using monitoring data. The monitor locations are often outside the control of researchers, and previous studies have shown that "preferential sampling" of monitoring locations can adversely affect exposure prediction and subsequent health effect estimation. We adopt a slightly different definition of preferential sampling than is typically seen in the literature, which we call population-based preferential sampling. Population-based preferential sampling occurs when the location of the monitors is dependent on the subject locations. We show the impact that population-based preferential sampling has on exposure prediction and health effect estimation using analytic results and a simulation study. A simple, one- parameter model is proposed to measure the degree to which monitors are preferentially sampled with respect to population density.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 01-Aug-2018

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 1-Published
Journal: Data Mining and Knowledge Discovery
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: Adjusting for scorekeeper bias in NBA box scores
Authors: Matthew van Bommel
Keywords: Statistics
Abstract: BoxscorestatisticsintheNationalBasketballAssociationareusedtomea- sure and evaluate player performance. Some of these statistics are subjective in nature and since box score statistics are recorded by scorekeepers hired by the home team for each game, there exists potential for inconsistency and bias. These inconsistencies can have far reaching consequences, particularly with the rise in popularity of daily fantasy sports. Using box score data, we estimate models able to quantify both the bias and the generosity of each scorekeeper for two of the most subjective statistics: assists and blocks. We then use optical player tracking data for the 2015–2016 season to improve the assist model by including other contextual spatio-temporal variables such as time of possession, player locations, and distance traveled. From this model, we present results measuring the impact of the scorekeeper and of the other contextual variables on the probability of a pass being recorded as an assist.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 4-Under Review
Journal: Structural Health Monitoring
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: Robust structural health monitoring under environmental and operational uncertainty
Authors: Anthony Liu, Lazhi Wang, Charles Farrar, Luke Bornn
Keywords: Statistics
Abstract: Existing methods for structural health monitoring are limited due to their sensitivity to changes in environmental and operational conditions, which can obscure the indications of damage by introducing nonlinearities and other types of noise into the structural response. In this paper, we introduce a novel approach using state-space probability models to infer the conditions underlying each time step, allowing the definition of a damage metric robust to environmental and operational variation. We define algorithms for training and prediction, describe how the algorithm can be applied in both the presence and absence of measurements for external conditions, and demonstrate the method’s performance on data acquired from a laboratory structure that simulates the effects of damage and EOV on bridges.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support: **Y**

Nothing to report in the uploaded pdf (see accomplishments)